Modular structure of the genetic code provides a link to the complete RNA code

Brian K. Davis

Research Foundation of Southern California, Inc.
P.O. Box 13595, La Jolla, CA 92039

Proteins are assembled almost exclusively from a set of only twenty amino acids that broadly separate into five biosynthesis families, originating from different precursors in the ancient pathways of central metabolism. Same-family amino acids generally charge phylogenetically- and structurally-related tRNA that read codons within a contiguous region of the genetic code, termed a code domain. Complementary anticodons are usually non-contiguous and pair tRNA of different code domains. They displayed elevated tRNA N2-base complementarity (reduced N2-base identity) at this G,C rich site, low pre-species-divergence sequence identity, low anticodon contiguity and negatively correlated amino acid path-distances. Sibling amino acid reliance on related tRNA, cognate with contiguous codons, links genetic code expansion with the growth of nascent tRNA-dependent amino acid synthesis pathways. Path-identity elements in these tRNA were credited with specifying amino acids in the RNA code at the origin of protein synthesis.